



This is a repository copy of *From bed to bench: how in silico medicine can help ageing research.*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/138566/>

Version: Published Version

Article:

Viceconti, M. orcid.org/0000-0002-2293-1530 and Dall'Ara, E.
orcid.org/0000-0003-1471-5077 (2018) From bed to bench: how in silico medicine can help ageing research. *Mechanisms of Ageing and Development*. ISSN 0047-6374

<https://doi.org/10.1016/j.mad.2018.07.001>

Reuse

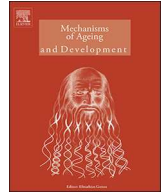
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



From bed to bench: How *in silico* medicine can help ageing research

Marco Viceconti^{a,b,*}, Enrico Dall'Ara^{a,c}

^a INSIGNEO Institute for in silico Medicine, United Kingdom

^b Department of Mechanical Engineering, United Kingdom

^c Department of Oncology and Human Metabolism, University of Sheffield, United Kingdom

ARTICLE INFO

Keywords:

Osteoporosis

Mouse

Human

Subject-specific modelling

ABSTRACT

Driven by the raising ethical concerns surrounding animal experimentation, there is a growing interest for non-animal methods, *in vitro* or *in silico* technologies that can be used to reduce, refine, and replace animal experimentation. In addition, animal experimentation is being critically revised in regard to its ability to predict clinical outcomes. In this manuscript we describe an initial exploration where a set of *in vivo* imaging based subject-specific technologies originally developed to predict the risk of femoral strength and hip fracture in osteoporotic patients, were adapted to assess the efficacy of bone drugs pre-clinically on mice. The CT2S technology we developed generates subject-specific models based on Computed Tomography that can separate fractured and non-fractured patients with an accuracy of 82%. When used in mouse experiments the use of *in vivo* imaging and modelling was found to improve the reproducibility of Bone Mineral Content measurements to a point where up to 63% less mice would be required to achieve the same statistical power of a conventional cross-sectional study. We also speculate about a possible approach where animal-specific and patient-specific models could be used to better translate the observation made on animal models into predictions of response in humans.

1. Introduction

There is a growing interest for *non-animal methods*, *in vitro* or *in silico* technologies that can be used to reduce, refine, and replace animal experimentation (see for example this review relative to vision research (Combes and Shah, 2016)); the primary motivation is the raise in our societies of ethical concerns with animal experimentation.

But there is another problem that drives the critical revision of animal experimentation: According to a recent BIO report¹, based on a database of nearly 10,000 regulatory phase transitions, the Likelihood of Approval from phase I (after the compound has successfully passed the pre-clinical assessment, mostly based on animal experimentation) ranges between 26.1% in haematology to 5.1% in oncology, with an average across all indications of 9.6%. That mean that in the best-case scenario animal models are able to spot all safety and efficacy issues associated with a new compound only one time out of four.

In parallel there is a growing enthusiasm around the use of subject-specific computer modelling and simulation to support the clinical decision. Developed in Europe under the flag name of Virtual Physiological Human (VPH), and elsewhere with that of Physiome Project, these technologies are capable of accurately predict the disease

progression and the effect of different treatments in individual patients, for a variety of conditions (Viceconti and Hunter, 2016).

Computational methods have been used in the last few years to investigate various aspects on skeletal ageing (Lambers et al., 2015; Patel et al., 2014; Razi et al., 2015a,b; Yang et al., 2017). In this manuscript we describe an initial exploration where a set of VPH technologies, originally developed to predict the risk of femoral strength and hip fracture in osteoporotic patients, were adapted to assess the efficacy of bone drugs pre-clinically on mice. We also propose an approach where animal-specific and patient-specific models could be used to better translate the observation made on animal models into predictions of response in humans.

2. CT2S: predicting the individual risk of hip fracture in osteoporotic patients

Osteoporosis is a silent disease, in the sense that the patient has no symptoms, until a minor trauma, with impact energy much lower than what is normally required, produces a bone fracture (hereinafter called fragility fracture). The whole skeleton is affected, but the most common fractures are at the wrist, the ankle, the spine, and at the hip. The latter,

* Corresponding author at: INSIGNEO Institute for in silico Medicine, University of Sheffield, Pam Liversidge Building, S13JD, Sheffield, United Kingdom.

E-mail address: m.viceconti@Sheffield.ac.uk (M. Viceconti).

¹ See report "Clinical Development Success Rates 2006-2015" by BIO, Biomedtracker and Amplion - <https://tinyurl.com/jhg52dr> - Last accessed 22 Feb 2018.

<https://doi.org/10.1016/j.mad.2018.07.001>

Received 27 March 2018; Received in revised form 19 June 2018; Accepted 9 July 2018

0047-6374/ © 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

because of the severity of the consequences, is of particular clinical relevance.

Pharmacological treatments exist that slow down the progression of the disease, reducing the incidence of fragility fractures of 40% or better (Sanderson et al., 2016). But because this is a long-term therapy, with considerable associated costs, and some potential side effects, it is essential to recognise who are the patients with a sufficiently high risk of fragility fractures to justify such pharmacological treatment.

It is not trivial to define the accuracy with which a predictor identifies the patients at higher risk. One possible approach is to build artificial cohorts where patients who experienced a fragility fracture before being treated, and then for each of them identify another patient who did not had any fragility fracture yet, but match the fractured patient in gender, age, weight, and height. One of such cohorts, hereinafter referred as the *Sheffield Cohort*, includes 50 postmenopausal women with a fragility hip fracture and 50 control cases (no fractures), pair-matched by gender, age, weight, and height (Yang et al., 2014). With such cohorts it is possible to compare the accuracy of risk predictors in term of their ability to discriminate the fractured patients and non-fractured ones; this will be referred to hereinafter as *discrimination accuracy*.

The risk of experiencing a fragility hip fracture is strongly related to the *side-fall strength* of the patient's femur, intended as the force required to fracture the femoral bone, under loading conditions similar to those observed during side-fall impact with the floor. Strength can be measured only invasively; however, a number of methods use medical imaging to acquire non-invasively information about shape and/or the degree of mineralisation of the patient's bone and use this information to estimate the strength. The current clinical standard involves measuring the areal Bone Mineral Density (aBMD) at the hip region using a specialised medical imaging technology called Dual X-ray Absorptiometry (DXA). Since 1985 (Basu et al., 1985), a number of research groups have refined the methods involved with an alternative approach, where quantitative computed tomography (QCT) of the femoral region is used to generate a subject-specific finite element model (SSFE) of the patient's bone, a computer model that is used to predict the femoral strength.

In a recent meta-analysis, conducted over a set of studies run by five different groups and that involved over 400 cadaver femurs, we showed that QCT-SSFE is more accurate (84–85%) than DXA-aBMD (77–81%) in predicting femoral strength in side fall, as measured in cadaver femurs (Viceconti et al., 2018). However, until recently, QCT-SSFE was reported to be not significantly more accurate than DXA-aBMD in predicting the risk of hip fracture (Zysset et al., 2015), due to a number of methodological issues. The CT2S technology developed in Sheffield has addressed most of these issues effectively. The femoral geometry is segmented from the CT images, and then automatically meshed with parabolic 10-node tetrahedral elements using a modified version of the advancing front method. The resulting mesh is projected back onto the CT images, and the calibrated CT attenuation coefficients are numerically averaged over the volume of each finite element, assuming trilinear interpolation between voxel values, in order to generate heterogeneous material properties mapping. As human femurs exhibit a clearly fragile failure under side fall loading, we use a maximum strain failure criterion with different limits for tension and compression. Over the years the method was refined through extensive methodological research (Polgar et al., 2001; Taddei et al., 2004; Viceconti and Taddei, 2003; Viceconti et al., 1999a,b; Zannoni et al., 1998), subjected to extensive verification (Taddei et al., 2006; Viceconti et al., 2004), and equally extensive validation against strain and strength measurements on cadaver femurs (Grassi et al., 2012; Schileo et al., 2014, 2008a; Schileo et al., 2008b, 2007). The method was further developed to be used as a clinical predictor of the risk of fracture (Falcinelli et al., 2014; Qasim et al., 2016); on the Sheffield cohort the stratification accuracy of CT2S is 82%, while that of DXA-aBMD is 75% (Viceconti et al., 2018).

3. DigiMouse: imaging & modelling technologies to quantify bone adaptation to therapies

The generic term of “bone drugs” indicates a variety of pharmaceutical products (Alendronate, Risedronate, Ibandronate, Zoledronic acid, Denosumab, Raloxifene, Teriparatide, Abaloparatide, etc.) which through fairly different mechanisms of actions slow down or reverse the progressive loss of bone mass. The most common animal models used to test the efficacy of new bone drugs use wild-type (WT) female mice of C57BL/6 or Balb-C strains, typically aged around 14 weeks, who all receive an abdominal surgery; for half this involves the removal of the ovaries (OVX), and for the other half only the sham surgery (SHAM). Each of the two groups is subdivided in two sub-groups, one that is injected with the drug being tested, the other with just the vehicle. After a period of treatment, typically four weeks, all animals are sacrificed, one bone (typically the tibia or the femur) is harvested, and two small portions of 2–3 cubic millimetres of bone tissue below the proximal growth plate and in the mid-shaft are scanned with micro-computed tomography (microCT) (Ruegsegger et al., 1996) in order to evaluate the properties of the cancellous and cortical bone, respectively (Bouxsein et al., 2010). From the obtained images 3D histomorphometric measurements can be performed to describe the tissue density and morphology. The most important parameter is called bone volume fraction (BV/TV), which is the volume of bone that is mineralised in the considered region of interest (ROI) over the total volume of the ROI. The higher is the BV/TV of the treated animals compared to those non-treated, the more effective is considered the treatment. The high-resolution microCT images allow also measuring other important microstructural parameters as trabecular thickness (Tb.Th), trabecular number (Tb.N), trabecular spacing (Tb.Sp), cortical area (Ct.A) and cortical thickness (Ct.Th), which provide an assessment of local bone quality (Bouxsein et al., 2010). For example, see (Johnston et al., 2007) who investigated the effect of a combined Teriparatide and Alendronate treatment with this animal model.

Although these are rarely explicitly stated, this model relies on a number of assumptions:

- 1 The ovariectomy produces in mice a systemic osteopenia similar to age-related osteoporosis in humans;
- 2 The sham surgery is required to normalise for the systemic effect of surgery;
- 3 The proximal tibia cancellous bone and the measurement of the properties of the cortical at a thin portion of the mid-shaft well represent global skeletal changes (negligible spatial gradient).
- 4 The changes at four weeks are representative of the long-term response (negligible temporal gradient after the time point);

The above-mentioned analyses can also be performed longitudinally by using an *in vivo* microCT imaging (Dall'Ara et al., 2016) where each mouse would act as its own control, instead of the standard cross-sectional design. In a first study (Lu et al., 2015) we explored surface-to-surface distance as a measure of bone remodelling between two consecutive time points, subsequent to proper image segmentation and Boolean operations, an approach similar to that used by Ralph Müller's group at ETH Zurich to investigate the bone remodelling of mouse caudal vertebrae (Christen and Muller, 2017; Schulte et al., 2011).

While the method was found very accurate by using repeated scans, when applied to longitudinal data it highlighted that even after 14 weeks of age mouse tibiae continue to grow, and even if the linear growth per week is small, in the order of a few hundred micrometres, it biases heavily the bone remodelling quantification, making difficult to separate changes due to growth from those due to bone remodelling. To compensate for this, a new approach was adopted (Lu et al., 2016). The 3D images obtained at different time-points were all aligned to the one generated at the first time-point (baseline) using rigid registration. In each image the tibia was segmented by thresholding and its length

Longitudinal assessment of bone properties

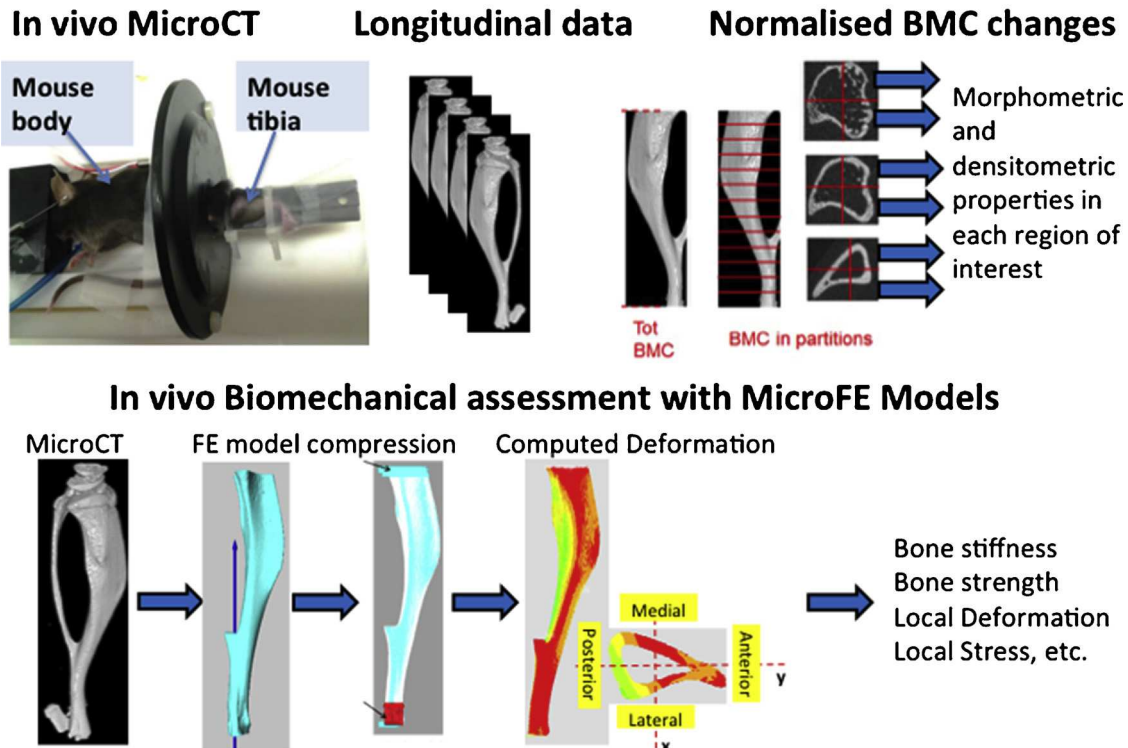


Fig. 1. Top: workflow to assess morphometric, densitometric and mechanical properties of the mouse tibia over time from longitudinal microCT images. The longitudinal microCT data are used to evaluate the change in total bone mineral content (Tot BMC) over time with respect to baseline. Similarly the normalised changes in bone mineral content (BMC) or other morphometric and densitometric properties are computed in 40 regions of the tibia by dividing it into ten longitudinal sections and by dividing each section in four quadrants. Bottom: each microCT image can then be converted into a micro finite element (microFE) model in order to estimate the changes in mechanical properties over time. For more details about the workflow please refer to (Lu et al., 2017).

Table 1

Theoretical sample size reduction (with 95% confidence interval, CI) if data from the longitudinal study are analysed by accounting for baseline measurements or if they are analysed as standard cross-sectional study. The data are reported for morphometric and densitometric parameters: Bone Volume Fraction (BV/TV), trabecular thickness (Tb.Th), trabecular number (Tb.N), cortical thickness (Ct.Th), total trabecular bone mineral content (Tb.BMC), total cortical bone mineral content (Ct.BMC) and total bone mineral content (Tot.BMC). The intra-class correlation (ICC) is also reported for each variable.

	Variable	ICC point estimate	Theoretical sample size reduction (95% CI)
Standard Morphometric parameters	BV/TV	0.59	34% (20% to 53%)
	Tb.Th	0.13	13% (6% to 28%)
	Ct.Th	0.64	28% (18% to 41%)
	Tb.N	0.27	7% (2% to 18%)
BMC parameters	Tb.BMC	0.71	51% (35% to 68%)
	Ct.BMC	0.87	63% (54% to 71%)
	Tot.BMC	0.87	63% (54% to 71%)

calculated. The analyses included only the 80% of the bone below the proximal growth plate, which was subdivided in ten equally spaced sections, which were further subdivided into four anatomical oriented (medial, lateral, anterior, posterior) compartments (Fig. 1, top). Using this anatomical partitioning, we evaluated the reproducibility of various measurements, finding that the total Bone Mineral Content (BMC) obtained by the necessary densitometric calibration was more reproducible than other morphometric indicators, such BV/TV. A simple power calculation showed that by using a longitudinal approach

enabled by *in vivo* imaging and the more reproducible BMC as metric of bone adaptation, we could reduce of 63% the number of animals required in each group to achieve the same statistical power, with respect to the traditional cross-sectional study described above (see Table 1). This reduction of the required animals for each group should be added to the several animals that do not need to be sacrificed at each time point. For example, if we are interested in measuring the bone changes at six different time points between two groups of 10 mice, with a standard cross section study we would need 120 mice. With a longitudinal approach we would just need 8 mice.

Nevertheless, one potential concern for the longitudinal assessment is that the X-ray ionising radiation that the repeated microCT scans involve (approximately 500 mGy per scan) might affect the bone remodelling process, altering the very same biological mechanism we want to observe. In small group of mice ($N = 5$ per group) treated with parathyroid hormone (PTH) or with vehicle, the right leg was scanned 8 times *in vivo* (once per week between weeks 14 and 22 of age, excluding week 15 that was skipped) and at week 24 both right (irradiated) and left (not irradiated) tibiae were scanned *ex vivo*. The differences between the total BMC of tibiae were between 3.7% and 8.0% (difference for control group: $5.8 \pm 1.9\%$, difference for treated group $5.1 \pm 0.9\%$) (Lu et al., 2017). This difference was slightly larger than that found between left and right tibiae of 10 control mice of the same age (22 weeks) that were not irradiated ($2.0 \pm 1.3\%$, range: 0.3% to 3.7%). Therefore, as caution principle, we recently revised the longitudinal protocol to perform the scans every two weeks instead than weekly and to half the scanning time which leads to a reduction of the nominal ionising radiation dose from 513 mGy to 256 mGy per scan. This reduction in scanning time was not associated with significant

differences in the estimated densitometric and mechanical properties of the tibia (Oliviero et al., 2017).

The new method described here opened additional possibilities, beyond the reduction of animal experimentation. A first important benefit is that for the first time it was possible to observe the effect of an intervention over a large anatomical site (the whole tibia), and across time. We used this to explore how the injection of PTH would manifest in space and time, with respect to WT mice without any intervention (Lu et al., 2017). Indeed, the study found that the anabolic effect of PTH follows a specific spatial-temporal pattern, confirming that spatial and temporal gradients are not negligible.

Furthermore, the longitudinal scanning of the whole mouse tibia allows to use mouse-specific computer modelling for the non-invasive prediction of the bone strength, a much more informative quantity to investigate the effect of an intervention on the mechanical competence of the skeleton. This application is explored in detail in the following section.

4. Predicting bone strength non-invasively in mice

While animal models are useful for the prediction of effect of novel interventions on the density, microstructure and mechanical competence of the bone tissue, the differences in anatomy, physiology and kinematics make it impossible to have a preclinical counterpart of the clinical estimation of the risk of fracture. However, the bone mechanical competence, an endpoint in the assessment of interventions which is much more clinically relevant than standard densitometric and morphometric measurements, should be measured or estimated when possible.

As described above, from longitudinal microCT images we can measure a large number of densitometric and morphometric parameters that can be affected by ageing, diseases and/or interventions. In most cases it is hard to quantify the efficacy of the interventions from this multi-dimensional dataset: it is hard to assess how good is a treatment that improves one of the morphometric parameters or only the bone density in a small region of the considered bone. In this context the measurement or estimation of a few mechanical parameters can provide a much easier interpretation of the effect of interventions, which target an improvement of the bone mechanical competence. These parameters include for the example the load required to induce a fracture (failure load or bone strength), the load required to start permanent deformation in the bone (yield load), the bone stiffness and the energy required to induce a fracture in the bone.

The mechanical properties of mice bones can be measured in the laboratory through standard mechanical testing (Schrieffer et al., 2005). However, due to the complex shape and the small size of these *ex vivo* analyses are affected by experimental errors (Wallace et al., 2014) or are limited to loading conditions, which are less physiologically relevant (e.g. three-point bending test). Moreover, the estimation of yield or failure loads require destructive mechanical tests and can be therefore used only in standard cross-sectional studies, which require large number of animals and resources.

Similarly, to subject-specific QCT-based approaches developed for clinical applications, microCT-based finite element models (microFE) can be used to non-invasively estimate the bone mechanical properties longitudinally (Borg et al., 2018; Isaksson et al., 2009; Javaheri et al., 2018; Lambers et al., 2012; Liu et al., 2012; Lu et al., 2012; Lynch et al., 2010; Tiwari et al., 2018; Tiwari and Prasad, 2017; Ulrich et al., 1997, 1998; Wernle et al., 2010). The workflow to create a microFE model from the medical image is similar to that developed for clinical CT based finite elements for humans (Fig. 1, bottom). The bone voxels of the rigidly registered microCT images of the mouse tibia collected *in vivo* at the different time points are separated from the background and then converted into hexahedral finite elements. This *in silico* approach allows us to simulate different types of loading, the simplest and most important for the tibia being uniaxial compression.

Unfortunately, conversely to the clinical counterpart, the validation of the predictions of microFE outputs is limited to a few studies. While microFE models have been found to well estimate structural stiffness and strength of trabecular bone specimens (Schwiedrzik et al., 2016; Wolfram et al., 2010) and to well estimate local displacement under compression for trabecular bone (Chen et al., 2017) and vertebral bodies (Costa et al., 2017), little has been done to validate models of the mouse tibia. A few studies have found good agreement between the predicted local strains in a region of the external surface of the mouse tibia and experimental measurements performed with strain gauges (Patel et al., 2014; Razi et al., 2015b) or digital image correlation (Carriero et al., 2014; Pereira et al., 2015). However, the accuracy of the predictions of internal strains or of the structural mechanical properties of the tibia was never challenged. In a recent study performed in our group (unpublished data) we combined mechanical testing, microCT imaging, digital volume correlation and computational modelling to evaluate the ability of microFE models of the mouse tibia to predicting local and structural properties (Oliviero et al., 2018). The results of this pilot study (mechanical tests performed on six tibiae) showed good predictions of local displacements when compared to digital volume correlation measurements (coefficient of determination of the linear regressions between predicted and measured displacements greater than 82%, intercepts close to 0 and slopes in the range: 0.69–0.95), an image processing technique that measures of the full deformation field given two microCT images of an undeformed and deformed bone (Dall'Ara et al., 2014). Furthermore, both structural bone stiffness (difference between predicted and measured stiffness equal to $14 \pm 11\%$) and bone strength (difference between predicted and measured stiffness equal to $9 \pm 9\%$) were well predicted.

By using these validated microFE models applied to longitudinal microCT images we have estimated the effect of an anabolic treatment (parathyroid hormone, PTH1-34, base for the FDA approved drug against osteoporosis Teriparatide) on the bone mechanical competence of female C57BL/6 mice. In particular, we have shown that the tibia stiffness and strength normalized by baseline measurements (week 14 of age) were significantly higher after two weeks of treatment (treatment started at week 18 of age) compared to the control group (Lu et al., 2017).

5. The future: using models to translate from mouse to human

For the particular application addressed in this paper we have described two methodologies that allow respectively in humans and in mice to quantify non-invasively the changes in bone mass distribution in a given bone, and the relative changes in mechanical strength (and thus of the risk of fracture) under representative loading conditions.

It might be interesting to speculate how these two technologies could be combined. As a guiding example we will imagine a study aimed to establish if the combination of PTH and physical therapy is more effective than PTH alone in reducing the risk of bone fracture (Gardinier et al., 2015).

First, we would observe longitudinally the changes in bone mass distribution in the tibia, with and without superposition of some loading regime (say for example a vibration from the ground in standing position). Longitudinal *in vivo* microCT would give us the evolution of bone mass distribution over time, and animal-specific finite element models would give us an estimate of the actual biomechanical stimulus delivered in each point of that mouse's tibia under the vibration exercise, so to build a clear correlation between the dose of PTH injected, the intensity of the biomechanical stimulus, the anatomical location, and bone adaptation response over time.

A second experiment would be done on osteoporosis patients, under very similar conditions. We would use clinical CT to follow-up longitudinally changes in the bone mass distribution in the tibia but also in the femur. We would build patient-specific finite element models to calculate the actual biomechanical stimulus delivered in each point of

that patient's tibia by the vibration exercise.

As a first step we would use the matching data at the tibia to build a “scaling function” that tell us how the bone remodelling response observed in the mouse for a given normalised dose of PTH and for a given normalised level of biomechanical stimulus would correlate with that observed in the human under the same normalised conditions. We would then use this scaling function to predict the remodelling in the human femur, using the mouse response in the tibia and the actual normalised level of biomechanical stimulus in the femur predicted by the patient-specific model. We would then validate this mouse-human scaling model by comparing the predicted bone remodelling with that observed in human in the femur from longitudinal clinical CT.

Assuming the predictive accuracy is sufficiently good, we could now test only in the mouse a different dose or loading regime, or another anabolic composition, and use the scaling model to predict how such regime would change the bone mass distribution in the femur of the cohort of patients with the initial data of those enrolled in the previous study. Then we would use the CT2S technology to estimate how in that cohort such changes in femoral bone mass would reduce the actual risk of fracture at the hip.

All this would of course not replace the clinical trial, but we would go the clinical trial with much more confidence on the efficacy of our new treatment of what we can possibly achieve today with the current animal experimentation approaches.

6. Discussion

In this paper we propose one possible scenario for the use of modelling and simulation to reduce, refine and replace animal experimentation, and to improve the ability of animal experiments to predict the outcome in humans. Of course, there are other more radical scenarios: for example, to use the animal experiments simply to inform the inputs of *in silico* clinical trial population model, where a virtual population is “virtually treated” with a new drug, which effect is modelled from experiments on mice. An even more radical approach would be to skip animal experimentation altogether and inform such *in silico* clinical trial population model with the results of experiments done *in vitro*, on cell cultures. All these scenarios are plausible, and in fact in at least one case, the regulators have already accepted *in silico* testing as a full replacement of animal experiments (Man et al., 2014). The approach we proposed here is somehow incremental with respect to the current methods based on animal experimentation, and as such they may be more rapidly adopted.

A fundamental assumption in such scenario is that a “scaling function” exist to quantitatively translate observations made on mice to humans. Of course, case by case the existence of such function must be demonstrated. Where no such function could be found, this would cast serious doubts on the usefulness of that animal experiment to predict the response in humans; in such case, modelling and simulation could still be useful, this time to replace entirely animal experimentation.

7. Conclusions

In this manuscript we described an initial exploration where a set of VPH technologies, originally developed to predict the risk of femoral strength and hip fracture in osteoporotic patients, were adapted to assess the efficacy of bone drugs pre-clinically on mice. We also speculated about a possible approach where animal-specific and patient-specific models could be used to better translate the observation made on animal models into predictions of response in humans.

We are confident that *in vivo* imaging and subject-specific modelling technologies like the ones described here can not only reduce, refine, and in some cases even partially replace animal experimentation, but also provide a much more rigorous reasoning framework through which to interpret clinically the results of animal experiments.

Conflict of interest

The authors declare that they do not have any financial or personal relationships with other people or organisations that could have inappropriately influenced this study.

Acknowledgements

The authors acknowledge the partial support from the CompBioMed Centre of Excellence (H2020-EINFRA-2015-1-675451), the COST Action BM1402 (MouseAGE, supported by COST, European Cooperation in Science and Technology), the EPSRC funded project MultiSim (Grant N. EP/K03877X/1), the UK NC3R funded projects (Grant NC/K000780/1 and NC/R001073/1); and the FP7 European projects CHIC (grant N. 600841). The authors declare that they do not have any financial or personal relationships with other people or organisations that could have inappropriately influenced this study. The authors acknowledge Dr Peter Laud for the help with the statistics, and Ms Sara Oliviero for working on some of the preclinical data.

References

- Basu, P.K., Beall, A.G., Simmons, D.J., Vannier, M., 1985. 3-D femoral stress analysis using CT scans and p-version FEM. *Biomater. Med. Devices Artif. Organs* 13, 163–186.
- Borg, S.A., Buckley, H., Owen, R., Marin, A.C., Lu, Y., Eyles, D., Lacroix, D., Reilly, G.C., Skerry, T.M., Bishop, N.J., 2018. Early life vitamin D depletion alters the postnatal response to skeletal loading in growing and mature bone. *PLoS One* 13, e0190675.
- Bouxsein, M.L., Boyd, S.K., Christiansen, B.A., Guldberg, R.E., Jepsen, K.J., Muller, R., 2010. Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *J. Bone Miner. Res.* 25, 1468–1486.
- Carriero, A., Abela, L., Pitsillides, A.A., Shefelbine, S.J., 2014. Ex vivo determination of bone tissue strains for an *in vivo* mouse tibial loading model. *J. Biomech.* 47, 2490–2497.
- Chen, Y., Dall'Ara, E., Sales, E., Manda, K., Wallace, R., Pankaj, P., Viceconti, M., 2017. Micro-CT based finite element models of cancellous bone predict accurately displacement once the boundary condition is well replicated: a validation study. *J. Mech. Behav. Biomed. Mater.* 65, 644–651.
- Christen, P., Muller, R., 2017. *In vivo* visualisation and quantification of bone resorption and bone formation from time-lapse imaging. *Curr. Osteoporos. Rep.* 15, 311–317.
- Combes, R.D., Shah, A.B., 2016. The use of *in vivo*, *ex vivo*, *in vitro*, computational models and volunteer studies in vision research and therapy, and their contribution to the Three Rs. *Altern. Lab. Anim.* 44, 187–238.
- Costa, M.C., Tozzi, G., Cristofolini, L., Danesi, V., Viceconti, M., Dall'Ara, E., 2017. Micro finite element models of the vertebral body: validation of local displacement predictions. *PLoS One* 12, e0180151.
- Dall'Ara, E., Barber, D., Viceconti, M., 2014. About the inevitable compromise between spatial resolution and accuracy of strain measurement for bone tissue: a 3D zero-strain study. *J. Biomech.* 47, 2956–2963.
- Dall'Ara, E., Boudiffa, M., Taylor, C., Schug, D., Fiegle, E., Kennerley, A.J., Damianou, C., Tozer, G.M., Kiessling, F., Muller, R., 2016. Longitudinal imaging of the ageing mouse. *Mechan. Ageing Dev.* 160, 93–116.
- Falcinelli, C., Schileo, E., Balistreri, L., Baruffaldi, F., Bordini, B., Viceconti, M., Albisinni, U., Ceccarelli, F., Milandri, L., Toni, A., Taddei, F., 2014. Multiple loading conditions analysis can improve the association between finite element bone strength estimates and proximal femur fractures: a preliminary study in elderly women. *Bone* 67, 71–80.
- Gardinier, J.D., Mohamed, F., Kohn, D.H., 2015. PTH signaling during exercise contributes to bone adaptation. *J. Bone Miner. Res.* 30, 1053–1063.
- Grassi, L., Schileo, E., Taddei, F., Zani, L., Juszczak, M., Cristofolini, L., Viceconti, M., 2012. Accuracy of finite element predictions in sideways load configurations for the proximal human femur. *J. Biomech.* 45, 394–399.
- Isaksson, H., Grongroft, I., Wilson, W., van Donkelaar, C.C., van Rietbergen, B., Tami, A., Huiskes, R., Ito, K., 2009. Remodeling of fracture callus in mice is consistent with mechanical loading and bone remodeling theory. *J. Orthop. Res.* 27, 664–672.
- Javaheri, B., Razi, H., Piles, M., de Souza, R., Chang, Y.M., Maric-Mur, I., Hopkinson, M., Lee, P.D., Pitsillides, A.A., 2018. Sexually dimorphic tibia shape is linked to natural osteoarthritis in STR/Ort mice. *Osteoarthritis Cartil.* 26, 807–817.
- Johnston, S., Andrews, S., Shen, V., Cosman, F., Lindsay, R., Dempster, D.W., Iida-Klein, A., 2007. The effects of combination of alendronate and human parathyroid hormone (1-34) on bone strength are synergistic in the lumbar vertebra and additive in the femur of C57BL/6J mice. *Endocrinology* 148, 4466–4474.
- Labbers, F.M., Kuhn, G., Schulte, F.A., Koch, K., Muller, R., 2012. Longitudinal assessment of *in vivo* bone dynamics in a mouse tail model of postmenopausal osteoporosis. *Calcif. Tissue Int.* 90, 108–119.
- Labbers, F.M., Kuhn, G., Weigt, C., Koch, K.M., Schulte, F.A., Muller, R., 2015. Bone adaptation to cyclic loading in murine caudal vertebrae is maintained with age and directly correlated to the local micromechanical environment. *J. Biomech.* 48, 1179–1187.
- Liu, X.S., Ardeshirpour, L., VanHouten, J.N., Shane, E., Wysolmerski, J.J., 2012. Site-

- specific changes in bone microarchitecture, mineralization, and stiffness during lactation and after weaning in mice. *J. Bone Miner. Res.* 27, 865–875.
- Lu, Y., Thiagarajan, G., Nicoletta, D.P., Johnson, M.L., 2012. Load/strain distribution between ulna and radius in the mouse forearm compression loading model. *Med. Eng. Phys.* 34, 350–356.
- Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., Viceconti, M., 2015. Evaluation of in-vivo measurement errors associated with micro-computed tomography scans by means of the bone surface distance approach. *Med. Eng. Phys.* 37, 1091–1097.
- Lu, Y., Boudiffa, M., Dall'Ara, E., Bellantuono, I., Viceconti, M., 2016. Development of a protocol to quantify local bone adaptation over space and time: quantification of reproducibility. *J. Biomech.* 49, 2095–2099.
- Lu, Y., Boudiffa, M., Dall'Ara, E., Liu, Y., Bellantuono, I., Viceconti, M., 2017. Longitudinal effects of parathyroid hormone treatment on morphological, densitometric and mechanical properties of mouse tibia. *J. Mech. Behav. Biomed. Mater.* 75, 244–251.
- Lynch, M.E., Main, R.P., Xu, Q., Walsh, D.J., Schaffler, M.B., Wright, T.M., van der Meulen, M.C., 2010. Cancellous bone adaptation to tibial compression is not sex dependent in growing mice. *J. Appl. Physiol.* (Bethesda, Md.: 1985) 109, 685–691.
- Man, C.D., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., Cobelli, C., 2014. The UVA/PADOVA type 1 diabetes simulator: New features. *J. Diabetes Sci. Technol.* 8, 26–34.
- Oliviero, S., Lu, Y., Viceconti, M., Dall'Ara, E., 2017. Effect of integration time on the morphometric, densitometric and mechanical properties of the mouse tibia. *J. Biomech.* 65, 203–211.
- Oliviero, S., Giorgi, M., Dall'Ara, E., 2018. Validation of finite element models of the mouse tibia using digital volume correlation. *J. Mech. Behav. Biomed. Mater.* 86, 172–184. <https://doi.org/10.1016/j.jmbbm.2018.06.022>. Jun 18.
- Patel, T.K., Brodt, M.D., Silva, M.J., 2014. Experimental and finite element analysis of strains induced by axial tibial compression in young-adult and old female C57BL/6 mice. *J. Biomech.* 47, 451–457.
- Pereira, A.F., Javaheri, B., Pitsillides, A.A., Shefelbine, S.J., 2015. Predicting cortical bone adaptation to axial loading in the mouse tibia. *J. R. Soc. Interface* 12, 0590.
- Polgar, K., Viceconti, M., O'Connor, J.J., 2001. A comparison between automatically generated linear and parabolic tetrahedra when used to mesh a human femur. *Proc. Inst. Mech. Eng. [H]* 215, 85–94.
- Qasim, M., Farinella, G., Zhang, J., Li, X., Yang, L., Eastell, R., Viceconti, M., 2016. Patient-specific finite element estimated femur strength as a predictor of the risk of hip fracture: the effect of methodological determinants. *Osteoporos. Int.* 27, 2815–2822.
- Razi, H., Birkhold, A.I., Weinkamer, R., Duda, G.N., Willie, B.M., Checa, S., 2015a. Aging leads to a dysregulation in mechanically driven bone formation and resorption. *J. Bone Miner. Res.* 30, 1864–1873.
- Razi, H., Birkhold, A.I., Zaslansky, P., Weinkamer, R., Duda, G.N., Willie, B.M., Checa, S., 2015b. Skeletal maturity leads to a reduction in the strain magnitudes induced within the bone: a murine tibia study. *Acta Biomater.* 13, 301–310.
- Rueggsegger, P., Koller, B., Muller, R., 1996. A microtomographic system for the non-destructive evaluation of bone architecture. *Calcified Tissue Int.* 58, 24–29.
- Sanderson, J., Martyn-St James, M., Stevens, J., Goka, E., Wong, R., Campbell, F., Selby, P., Gittos, N., Davis, S., 2016. Clinical effectiveness of bisphosphonates for the prevention of fragility fractures: a systematic review and network meta-analysis. *Bone* 89, 52–58.
- Schileo, E., Taddei, F., Malandrino, A., Cristofolini, L., Viceconti, M., 2007. Subject-specific finite element models can accurately predict strain levels in long bones. *J. Biomech.* 40, 2982–2989.
- Schileo, E., Dall'ara, E., Taddei, F., Malandrino, A., Schotkamp, T., Baleani, M., Viceconti, M., 2008a. An accurate estimation of bone density improves the accuracy of subject-specific finite element models. *J. Biomech.* 41, 2483–2491.
- Schileo, E., Taddei, F., Cristofolini, L., Viceconti, M., 2008b. Subject-specific finite element models implementing a maximum principal strain criterion are able to estimate failure risk and fracture location on human femurs tested in vitro. *J. Biomech.* 41, 356–367.
- Schileo, E., Balistreri, L., Grassi, L., Cristofolini, L., Taddei, F., 2014. To what extent can linear finite element models of human femora predict failure under stance and fall loading configurations? *J. Biomech.* 47, 3531–3538.
- Schrieffer, J.L., Robling, A.G., Warden, S.J., Fournier, A.J., Mason, J.J., Turner, C.H., 2005. A comparison of mechanical properties derived from multiple skeletal sites in mice. *J. Biomech.* 38, 467–475.
- Schulte, F.A., Lambers, F.M., Kuhn, G., Muller, R., 2011. In vivo micro-computed tomography allows direct three-dimensional quantification of both bone formation and bone resorption parameters using time-lapsed imaging. *Bone* 48, 433–442.
- Schwiedrzik, J., Gross, T., Bina, M., Pretterklieber, M., Zysset, P., Pahr, D., 2016. Experimental validation of a nonlinear muFE model based on cohesive-frictional plasticity for trabecular bone. *Int. J. Numer. Methods Biomed. Eng.* 32, e02739.
- Taddei, F., Pancanti, A., Viceconti, M., 2004. An improved method for the automatic mapping of computed tomography numbers onto finite element models. *Med. Eng. Phys.* 26, 61–69.
- Taddei, F., Martelli, S., Reggiani, B., Cristofolini, L., Viceconti, M., 2006. Finite-element modeling of bones from CT data: sensitivity to geometry and material uncertainties. *IEEE Trans. Bio-Med. Eng.* 53, 2194–2200.
- Tiwari, A.K., Prasad, J., 2017. Computer modelling of bone's adaptation: the role of normal strain, shear strain and fluid flow. *Biomechan. Model. Mechanobiol.* 16, 395–410.
- Tiwari, A.K., Kumar, R., Tripathi, D., Badhyal, S., 2018. In silico modeling of bone adaptation to rest-inserted loading: strain energy density versus fluid flow as stimulus. *J. Theoret. Biol.* 446, 110–127.
- Ulrich, D., Hildebrand, T., Van Rietbergen, B., Muller, R., Rueggsegger, P., 1997. The quality of trabecular bone evaluated with micro-computed tomography, FEA and mechanical testing. *Stud. Health Technol. Inform.* 40, 97–112.
- Ulrich, D., van Rietbergen, B., Weinans, H., Rueggsegger, P., 1998. Finite element analysis of trabecular bone structure: a comparison of image-based meshing techniques. *J. Biomech.* 31, 1187–1192.
- Viceconti, M., Hunter, P., 2016. The virtual physiological human: Ten years after. *Annu. Rev. Biomed. Eng.* 18, 103–123.
- Viceconti, M., Taddei, F., 2003. Automatic generation of finite element meshes from computed tomography data. *Crit. Rev. Biomed. Eng.* 31, 27–72.
- Viceconti, M., Zannoni, C., Testi, D., Cappello, A., 1999a. CT data sets surface extraction for biomechanical modeling of long bones. *Comput. Methods Programs Biomed.* 59, 159–166.
- Viceconti, M., Zannoni, C., Testi, D., Cappello, A., 1999b. A new method for the automatic mesh generation of bone segments from CT data. *J. Med. Eng. Technol.* 23, 77–81.
- Viceconti, M., Davinelli, M., Taddei, F., Cappello, A., 2004. Automatic generation of accurate subject-specific bone finite element models to be used in clinical studies. *J. Biomech.* 37, 1597–1605.
- Viceconti, M., Qasim, M., Bhattacharya, P., Li, X., 2018. Are CT-based finite element model predictions of femoral bone strength clinically useful? *Curr. Osteoporos. Rep.* 16 (June), 216–223.
- Wallace, R.J., Pankaj, P., Simpson, A.H., 2014. Major source of error when calculating bone mechanical properties. *J. Bone Miner. Res.* 29, 2697.
- Wernle, J.D., Damron, T.A., Allen, M.J., Mann, K.A., 2010. Local irradiation alters bone morphology and increases bone fragility in a mouse model. *J. Biomech.* 43, 2738–2746.
- Wolfram, U., Wilke, H.J., Zysset, P.K., 2010. Valid micro finite element models of vertebral trabecular bone can be obtained using tissue properties measured with nanoindentation under wet conditions. *J. Biomech.* 43, 1731–1737.
- Yang, L., Udall, W.J., McCloskey, E.V., Eastell, R., 2014. Distribution of bone density and cortical thickness in the proximal femur and their association with hip fracture in postmenopausal women: a quantitative computed tomography study. *Osteoporos. Int.* 25, 251–263.
- Yang, H., Albiol, L., Chan, W.L., Wulsten, D., Seliger, A., Thelen, M., Thiele, T., Spevak, L., Boskey, A., Kornak, U., Checa, S., Willie, B.M., 2017. Examining tissue composition, whole-bone morphology and mechanical behavior of GoraB(Prx1) mice tibiae: a mouse model of premature aging. *J. Biomech.* 65, 145–153.
- Zannoni, C., Mantovani, R., Viceconti, M., 1998. Material properties assignment to finite element models of bone structures: a new method. *Med. Eng. Phys.* 20, 735–740.
- Zysset, P., Qin, L., Lang, T., Khosla, S., Leslie, W.D., Shepherd, J.A., Schousboe, J.T., Engelke, K., 2015. Clinical use of quantitative computed tomography-based finite element analysis of the hip and spine in the management of osteoporosis in adults: the 2015 ISCD official positions-part II. *J. Clin. Densitom.* 18, 359–392.